INTERNATIONAL JOURNAL OF CIRCUMPOLAR HEALTH
2020, VOL. 79, 1764284
https://doi.org/10.1080/22423982.2020.1764284

Taylor & Francis
Taylor & Francis Group

Methodological issues on agreement between self-reported and central cancer registry-recorded prevalence of cancer in the Alaska EARTH study

Mehdi Naderi [a] and Siamak Sabour[b,c]

[a]Clinical Research Development Centre, Taleghani and Imam Ali Hospital, Kermanshah University of Medical Sciences, Kermanshah, I.R. Iran; [b]Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran; [c]Safety Promotions and Injury Prevention Research Centre, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

ARTICLE HISTORY Received 13 August 2019; Accepted 29 April 2020

KEYWORDS Alaska Native; tumour registry; study cohort; self-report; health literacy; reliability; methodological issues

Dear Editor,

We read with interest the article of Nash SH et al., published in the Int J Circumpolar Health 2019 Dec [1]. Determination of the agreement between self-reported and registry-recorded site-specific cancer diagnoses in a cohort of Alaska Native people [1]. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were used to calculate the agreement between self-reported and registry-recorded cancer diagnosis and finally, kappa values were calculated to differentiate between true agreement and agreement that may be expected due to chance [1]. Based on the results of their study for all-sites, as well as each common site, specificity was more than 98% for all cancer sites, whereas sensitivity, PPV and NPV for (colorectal cancer and prostate cancer), (colorectal cancer and female breast cancers) and all cancer sites were (78.6% and 100.0%), (52.4% and 84.8%) and more than 99.6%. Kappa values also varied by cancer site: values were high for female breast and prostate cancers ($\kappa = 0.86$ for both sites), and moderate for colorectal cancer ($\kappa = 0.63$). The agreement measures in strata of demographic characteristic were as follows: for cancer (all-sites), sensitivity was greater among males, those aged 18–50 years at study enrolment, those living in an urban area and those who spoke English as their primary language at home. Neither specificity nor NPV varied substantially by demographic characteristic. In contrast, higher PPV was observed among males, those aged 50+ years at study enrolment, those residing in an urban area and those reporting non-English or both as the primary language(s) spoken at home. The pattern was similar for kappa, where we observed greater values among males, those aged 50 + years at study enrolment and those residing in an urban area [1].

Reliability and validity are two completely different methodological issues. Sensitivity, specificity, (PPV), (NPV), likelihood ratios positive and negative (LR+ & LR-) are among the estimates to assess validity of a diagnostic test and have nothing to do with reliability [2,3]. The amount of kappa used to calculate reliability of qualitative and rank variables has some drawbacks that we describe below: The first problem is that the kappa value is strongly dependent on the prevalence and number of categories. Finally, another problem occurs when the two voters differ in the marginal distribution of their responses [2,4–6]. Table 1 illustrates these problems with a hypothetical example that ultimately shows the kappa value with the prevalence and the number of categories with different values (0.44 as moderate and 0.80 as very good).

The authors came to the conclusion that the good agreement is between self-reported and registry-recorded cancer history that may be the result of the high quality of care within the Alaska Tribal Health System [1]. Such a conclusion may be due to

Table 1. The kappa and weighted kappa values for calculating agreement between 2 raters for more than 2 categories.

		Raters 1			Sum
	Grade	**1**	**2**	**3**	
Raters 2	**1**	60	20	1	81
	2	2	12	4	18
	3	3	11	11	25
Sum		65	43	16	124
		Estimate			
Kappa		0.43			
Weighted kappa		0.63			

CONTACT Siamak Sabour s.sabour@sbmu.ac.ir Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Chamran Highway, Velenjak, Daneshjoo Blvd, Tehran, I.R. Iran

inappropriate use of the statistical test, which ultimately leads to a misleading message.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

ORCID

Mehdi Naderi http://orcid.org/0000-0002-5608-6582

References

[1] Nash SH, Day G, Hiratsuka VY, et al. Agreement between self-reported and central cancer registry-recorded prevalence of cancer in the Alaska EARTH study. Int J Circumpolar Health. 2019 Dec;78(1):1571383. PubMed PMID: 30724720.

[2] Sabour S, Ghassemi F. The validity and reliability of a signal impact assessment tool: statistical issue to avoid misinterpretation. Pharmacoepidemiol Drug Saf. 2016 Oct;25(10):1215–1216. PubMed PMID: 27696610.

[3] Szklo M, Nieto FJ. Epidemiology beyond the basics. 3rd ed. Manhattan, New York: Jones and Bartlett Publisher; 2014.

[4] Sabour S, Dastjerdi EV. Reliability of four different computerized cephalometric analysis programs: a methodological error. Eur J Orthod. 2013 Dec;35 (6):848. PubMed PMID: 24132404.

[5] Naderi M, Sabour S. Reproducibility of diagnostic criteria associated with atypical breast cytology: a methodological issue. Cytopathology. 2018 Aug;29(4):396. PubMed PMID: 29688593.

[6] Sabour S. Reliability of immunocytochemistry and fluorescence in situ hybridization on fine-needle aspiration cytology samples of breast cancers: methodological issues. Diagn Cytopathol. 2016 Dec;44(12):1128–1129. PubMed PMID: 27491351.